# SKOLAR MD: A Model for Self-Directed, In-Context Continuing Medical Education

**Howard R. Strasberg MD MS, Thomas C. Rindfleisch MS[§], Steven Hardy PhD**
SKOLAR, part of Wolters Kluwer Health, Inc., Palo Alto, CA
[§]Stanford University School of Medicine, Palo Alto, CA

**ABSTRACT**

INTRODUCTION: SKOLAR has implemented a web-based CME program with which physicians can earn AMA Category 1 credit for self-directed learning. METHODS: Physicians researched their questions in SKOLAR and applied for CME. Physician auditors reviewed all requests across two phases of the project. A selection rule set was derived from phase one and used in phase two to flag a subset of requests for detailed review. The selection rule set is described. RESULTS: In phase one, SKOLAR received 1039 CME applications. Applicants frequently found their answer (94%) and would apply it clinically (93%). A linear regression analysis comparing time awarded to time requested (capped at actual time spent) had $R^2$=0.79. DISCUSSION: We believe that that this self-directed approach to CME is effective and an important complement to traditional CME programs. However, selective audit of self-directed CME requests is necessary to ensure validity of credits awarded.

## INTRODUCTION

It is well documented that practicing physicians have serious unmet information needs. Questions arise in as many as one in three patient visits and at least half of these questions are not answered even though answers frequently exist (~55-70%) in published resources[1 2 3 4].

Unfortunately, traditional CME has had relatively little success in meeting physicians' long-term educational needs[5]. Often structured around conferences or fixed lessons, traditional CME can become divorced from medical practice and has been found to have little or no impact on physician performance[6]. The effectiveness of traditional CME is inherently compromised because experts choose topics that they presume, but are not certain, should be important to the intended audience[7].

The most effective CME methods have used practice-enabling or reinforcing methods[8 9 10]. A "practice-learning" model of CME has been proposed[7] where CME becomes a means of improving patient outcomes through enhanced physician performance and is no longer an activity that is separate and discrete from patient care. SKOLAR, which grew out of the SHINE system at Stanford University[11], builds on this model.

SKOLAR simultaneously searches multiple knowledge resources using a single integrated search. The user's query is mapped automatically into the search syntax of each available resource, and is sent concurrently to all selected resources. The available resources include textbooks (e.g., Harrison's), drug information (e.g., Lexi-Comp), evidence-based medicine (e.g., Cochrane), PubMed, patient information, and organization-specific content such as local guidelines.

Search results are organized in the manner most appropriate for the particular resource. For example, hits in a textbook are presented within a table-of-contents structure that greatly facilitates navigation and search result selection.

SKOLAR keeps a record of each user's search activity including date and time, query terms, documents reviewed, and links followed. This log is a key element in the awarding of CME credits.

As part of a pilot project with the American Medical Association (AMA), physicians can earn Category 1 CME credit for legitimate self-initiated learning conducted in SKOLAR. An integral part of SKOLAR is a subsystem for awarding this CME credit based on learning about real patient-related questions.

## METHODS

We implemented a self-directed, in-context CME system in two phases. In both phases, physician auditors reviewed all CME requests. A key goal was to determine whether it would eventually be possible to move away from having to manually audit every single CME request. In phase two,

therefore, we implemented changes to focus audit reviews on dubious requests, based on the results from phase one.

Physicians were potentially eligible for CME credit for all the time spent in any SKOLAR session in which they identified at least one relevant document. The rationale for this policy was that additional and complementary reading should count as legitimate learning, even if it was not directly answering the question posed. Any given SKOLAR session could be used only once to apply for CME credit. Credits could be earned in any increments, but certificates were issued only in 15-minute increments. Credit was granted through the AMA and certificates were mailed to physicians quarterly. Each certificate reflected the cumulative CME credits earned by a given physician in a given quarter.

Physicians did the following to earn CME:

1. Identified an information need based on a clinical question arising from a patient encounter.
2. Searched and reviewed multiple SKOLAR knowledge sources until satisfied that an answer to the question had been found (this step may have been done in a single session or spread over multiple days)
3. Initiated a request for CME, thus:
    a. Typed in the question researched
    b. Stated whether the answer was found in SKOLAR
    c. Stated whether the answer was applicable to a clinical decision
    d. Typed in either the answer to the question OR if no answer was found what was nevertheless learned during the search
4. Reviewed personal SKOLAR logs to identify those documents most relevant to the learning on the topic
5. Estimated the amount of time spent learning in the system

In phase one, the system computed the time spent on each document for the auditors, but the physician users had to estimate personally the time they spent learning.

Each CME request was individually reviewed by one of eight physician auditors (i.e. peer reviewers) who examined the requests for evidence of gross discrepancies between the credit requested and the time actually spent or between the time actually spent and the time that might reasonably be spent reading the given set of documents.

Auditors also looked for clinical questions and written answers that made sense, reviewed the documents used for learning (particularly those identified by the learner as being most relevant to answering their clinical questions) and optionally provided comments back to the CME applicant. In phase one auditors used their judgment as to whether credit should be granted for reviewing documents unrelated to the user's clinical question.

Auditors had the following options: grant credit for the exact time requested, grant credit for more or less time than requested based on the actual time spent in the system, or (rarely) deny the request altogether. Auditors were instructed to seek a second opinion from an auditor colleague if they had any doubt as to how to handle a given CME request. Learners always had the opportunity to send feedback to auditors through an email address that kept the identity of the auditors hidden.

Phase one was scheduled for completion when 400 CME requests had been audited. The results from phase one were analyzed with both descriptive statistics and linear regression. Linear regression was used to compare the credit granted by the auditor to the credit requested by the physician user. The results from phase two, which is still in progress, were analyzed with descriptive statistics.

## RESULTS

### Phase One

Phase one ran from October 10, 2000 to February 6, 2002, during which time 1040 CME requests were received from 393 applicants. (Data analysis resources were not available prior to February 2002 so phase one ended up including a significantly larger sample than had been planned.)

One outlier was excluded, leaving 1039 CME requests for this analysis. Five auditors audited 99% of the requests, and three temporary auditors completed the remaining 1%. The 393 individuals included 109 (28%) in Family Practice, 107 (27%) in Internal Medicine, 46 (12%) in Pediatrics, 24 (6%) in Internal Medicine subspecialties, 24 (6%) in Otolaryngology, 21 (5%) in Allergy/

Immunology, 11 (3%) in general practice, 11 (3%) unknown and the rest in other specialties.

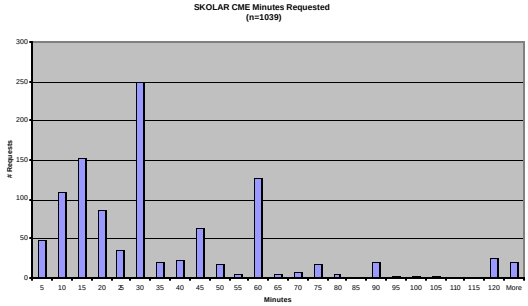


**Figure 1. Distribution of CME minutes requested.**

The average CME credit requested per application was 37 minutes (standard deviation 31 minutes). The distribution of minutes requested is presented in Figure 1.

On 94% of applications the applicant indicated that they found the answer to their clinical question in SKOLAR. On 93% of applications the applicant indicated that they would apply what they learned in SKOLAR to a clinical decision.

Table 1 provides a sample of the questions asked, and Table 2 provides a sample of an answer recorded by the physician.

1. What are the clinical and laboratory features of microscopic angiitis?
2. What is the current status of c-reactive protein testing in cardiology?
3. What are causes, symptoms, and treatments of the serotonin syndrome?
4. Is risperidone helpful in the treatment of stuttering?
5. What is pemphigus and how is it diagnosed and treated?
6. How are autothyroid antibodies and chronic urticaria related?
7. What is the long-term outcome for lyme disease peripheral neuropathy?
8. What is the current treatment for irritable bowel syndrome?

**Table 1 Sample of questions asked in the Stanford SKOLAR, M.D. CME system.**

In 49% of applications the physician applied for more time credit than the time logged by the system. In 21% of applications the physician applied for more than twice as much time credit than the time logged by the system.

Question: What are the clinical and laboratory features of microscopic angiitis?

Answer: Microscopic polyangiitis is a small vessel vasculitis with pathogenic, laboratory and clinical features similar to Wegener's granulomatosis. Characterized by ANCA positivity and variable involvement of renal and pulmonary systems. Important prognostic feature is whether or not pulmonary hemorrhage is noted; those with hemorrhage do less well. Treatment is generally with Cytoxan or some alkylating agent plus steroids

**Table 2 Sample answer provided by a physician requesting CME (edited for spelling errors only).**

A simple linear model relating minutes awarded and minutes requested had $R^2$=0.43 with significant scatter. We found that a better model related minutes awarded to minutes requested, but capping the minutes requested at the number of minutes the system actually logged. Figure 2 illustrates this latter model, which was a much better fit, having $R^2$=0.79.

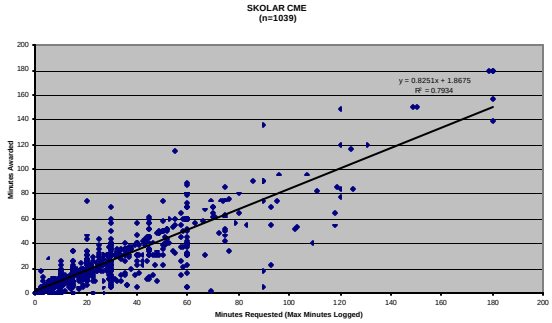


**Figure 2. Linear regression analysis relating minutes awarded by the auditor to minutes requested by the physician user, capping minutes requested at minutes actually logged.**

During phase one, auditors met with each other regularly to discuss the review process. Over this period of time, it became clear that it was difficult to judge objectively whether any given document was related to a user's clinical question, and therefore, this aspect of the audit process became gradually deemphasized. Similarly, at times some auditors felt that they disagreed with the answer that the physician user had reached. At a group meeting held on this topic, the auditors decided that since no one could be an expert in all areas of medicine, and since many medical questions are controversial, the auditors would neither comment on nor grant

credit on the basis of the apparent correctness of any given answer. This position also made sense from a liability perspective.

**Phase Two**

In phase two we modified the process in several ways, based on the results from phase one. In phase two physicians could not apply for more time than the system had actually logged. We decided to display this computed time for physicians at the time of their application, asking them then to estimate how much of that time they had actually spent studying.

Phase one results showed a good relationship between time granted and minutes requested (with a cap at minutes logged), so it seemed reasonable to differentiate audit requests in phase two. All requests were still routed for manual auditing but requests that raised certain 'red flags' were to be particularly carefully reviewed. Other requests were examined briefly for face validity and could then be approved with a single click (i.e. fast tracked). Auditors always read the question and answer and always had the option of doing a detailed review of any request that failed a face validity check.

The red flags were based primarily on excessive time spent reading a set of documents compared to the time spent in the past by other SKOLAR users reading similar documents. Also, all requests of two hours or more were flagged and auditors could flag users who had a history of seeking (intentionally or unintentionally) unreasonable credit for the documents read.

The modifications described above were developed and put in production on August 1, 2002. Phase two has no planned end date, but the analysis below reflects data through February 23, 2003. During phase two, SKOLAR received 1026 CME requests from 263 applicants. Two auditors audited and/or fast tracked all of these requests, with the following results:

| | Auditor reduced credit | Auditor granted requested credit |
|---|---|---|
| Request was 'red flagged' | 144 | 129 |
| Request marked eligible for 'fast track' | 53 | 700 |

Applying the flags retroactively to the phase one requests yielded a sensitivity of 24% and a specificity of 90%, assuming a positive gold standard was a request for which the auditor had reduced the time requested by any amount.

## DISCUSSION

In this pilot study we have implemented a system for CME in which physicians earn Category 1 credit for legitimate learning around real patient-related information needs. The efficient retrieval of knowledge to address these information needs is made possible through the use of an integrated search system (SKOLAR MD). The need to bring CME programs into the normal flow of patient care has been underscored by several years of literature describing the inefficacy of traditional CME models, and particularly, of didactic CME sessions[5-10].

We found that the vast majority of CME applicants indicated that they were able to find the answer to their clinical question in SKOLAR, and moreover, that they would indeed apply what they learned to a clinical decision. While these numbers do come from self-reporting, they speak very favorably to the value and efficacy of in-context CME.

The questions asked by applicants were intelligent and reflected a breadth of information needs consistent with previous work done by our group[12]. Interestingly enough, very few of those questions would likely be central themes in formal courses. On the other hand, our experience with in-context CME suggests that having a trusted, high quality integrated information resource at the point of care allows physicians to apply their learning immediately to clinical decision-making.

Our findings do not support the notion that in a self-directed learning system, physicians will always accurately report the time spent learning. In phase one we found that physicians overestimated the time spent learning relative to the actual time logged in almost half the cases. In phase two we found that the prevalence of CME requests for amounts of time not standing up to peer review was 19% ((144+53) out of 1026). We acknowledge that many explanations might exist as to why physician claims of time spent were sometimes inflated. Some possible explanations are: (1) physicians made time estimates in good faith but nevertheless their estimates were high; (2) some physicians may have taken more time than others to conceptualize the new information they learned; (3) in phase two, physicians took the course of

least effort and applied for the maximum time eligible, assuming that SKOLAR would adjust the credit if necessary and (4) in some cases, we believe rarely, physicians may have made deliberate attempts to request more credit than would seem reasonable to peer reviewers. We believe therefore, that some kind of system of checks and balances, guarding against intentional or unintentional inflation of credit, is required to maintain the integrity of a credit-awarding self-directed learning system.

To see widespread adoption of a learning system such as the one described requires that the system be practical. For this reason it is important to automate as many parts of it as possible, as we have started to do with the phase-two fast-track audits. We believe that the poor sensitivity found in the retrospective application of the red flags to the phase one data was more a reflection on the evolving gold standard than on the test itself. The gold standard changed in at least two ways. First, in phase one, auditors considered the relevance of individual documents read to the question asked, but in phase two, only overall document *set* relevance was a consideration. Second, in phase one, auditors considered the time spent reading each individual document, whereas in phase two, auditors concentrated more on the time spent reading the *set* of documents and less on the time spent reading each individual document. Our belief that the gold standard evolved is supported in part by the finding in phase two that, assuming that face validity checks were sufficient to detect significant discrepancies, the sensitivity of the red flags as a screening test increased from 24% to 73% (144/(144+53)). We are however continuing to refine the rules governing the selection of audits for detailed review.

We believe that the system described herein complements traditional CME and represents a viable model for CME today and in the future. Further studies are needed to determine whether adoption of this CME model results in improved patient care, the ultimate goal.

**ACKNOWLEDGMENTS**

It is with a mixture of gratitude and sadness that the authors acknowledge the seminal contributions of Kenneth L. Melmon, MD (deceased) to this work. The SKOLAR system and its use to promote self-initiated, in-context CME were Dr. Melmon's brainchild and highest priority research activity at the time of his death. This work has benefited in very many ways from his leadership, wisdom, and zeal.

**REFERENCES**

[1] Ely JW, Osheroff JA, Ebell MH et al. Analysis of questions asked by family doctors regarding patient care. BMJ 1999 Aug 7;319(7206):358-61.

[2] Chambliss ML, Conley J. Answering clinical questions. J Fam Pract 1996 Aug;43(2):140-4.

[3] Gorman P. Information needs in primary care: a survey of rural and nonrural primary care physicians. Medinfo 2001;10(Pt 1):338-42.

[4] Covell DG, Uman GC, Manning PR. Information needs in office practice: are they being met? Ann Intern Med 1985;103(4):596-9.

[5] Davis DA, Thomson MA, Oxman AD, Haynes RB. Changing physician performance. A systematic review of the effect of continuing medical education strategies [see comments]. JAMA 1995;274(9):700-5.

[6] Davis D, O'Brien MA, Freemantle N, Wolf FM, Mazmanian P, Taylor-Vaisey A. Impact of formal continuing medical education: do conferences, workshops, rounds, and other traditional continuing education activities change physician behavior or health care outcomes? JAMA 1999 Sep 1;282(9):867-74

[7] Barnes BE. Creating the practice-learning environment: using information technology to support a new model of continuing medical education. Acad Med 1998;73(3):278-81.

[8] Davis DA, Thomson MA, Oxman AD, Haynes RB. Evidence for the effectiveness of CME. A review of 50 randomized controlled trials [see comments]. JAMA 1992;268(9):1111-7.

[9] Thomson O'Brien MA, Oxman AD, Davis DA, Haynes RB, Freemantle N, Harvey EL. Audit and feedback: effects on professional practice and health care outcomes. Cochrane Database Syst Rev 2000;(2):CD000259.

[10] Leist JC, Green JS. Congress 2000: a continuing medical education summit with implications for the future. J Contin Educ Health Prof 2000 Fall;20(4):247-51.

[11] Hubbs PR, Tsai M, Dev P et al. The Stanford Health Information Network for Education: integrated information for decision making and learning. Proc AMIA Ann Fall Symp 1997:505-8.

[12] Strasberg HR, Hubbs PR, Rindfleisch TC, Melmon KL. Analysis of information needs of users of the Stanford Health Information Network for Education. Proc AMIA Symp 1999;:965-9